Research Management Review, Volume 18, Number 2
Fall/Winter 2011

Behind Door #3: The Hard-to-Please Grant Reviewer

Jeremy T. Miner
University of Wisconsin-Eau Claire

ABSTRACT

After months of waiting, the grant reviews came back: "excellent," "excellent," and "fair." What?! How can this be? Why is the third review so out of line with the first two?

On more than one occasion a principal investigator (PI) has been frustrated not only by a negative funding decision but more so by the accompanying reviewer evaluation forms that contain diametrically opposed feedback. Research administrators are in a prime position to help faculty unpack reviewer comments and to better understand the existence and nature of the hard-to-please grant reviewer. This article applies an analytic model to compare and contrast fourteen dimensions of grant structures and processes at the National Institutes of Health, National Science Foundation, and U.S. Department of Education that have an effect on grant outcomes. It examines five specific aspects of inconsistencies in ratings among grant reviewers that are of greatest concern to unsuccessful applicants. It also offers strategies that can be used by research administrators to propel the PI forward—to revise and resubmit, to pursue a different sponsor, or to consider a new line of investigation. PIs who anticipate the needs of reviewers and satisfy those expectations in the proposal increase their odds for grant success.

INTRODUCTION

On more than one occasion a research administrator has met with a principal investigator (PI) whose grant reviews have just come back and who was frustrated with the result. In some instances, the PI pushes the reviews across the table and demands to know, "How can this be?" Scanning the documents, the research administrator quickly identifies the point of concern:

Reviewer #1: Rating = *Excellent*
This proposed work addresses a problem of national relevance for students.

Reviewer #2: Rating = *Excellent*
Potentially great merit in the development of methods for teaching fundamental entry level math for STEM.

Reviewer #3: Rating = *Fair*
The intellectual merit of this proposal is rather low. The proposers understand that there is a problem, but they demonstrate no new insights into solutions.

In short, the PI expects the research administrator to be able to explain why the third review is so out of line with the first two. This article explores the existence and nature of the hard-to-please grant reviewer so that research administrators can better serve their faculty.

ANALYTIC FRAMEWORK

Research administrators are in a prime position to help faculty unpack reviewer comments, assuming that PIs do indeed wish to understand and learn from the written feedback. Some PIs may be more interested in venting or grieving. They could have concerns that this grant rejection will adversely impact their project plans, individual status, and career trajectory. PIs are used to being the experts; they might be less comfortable with anonymous challenges to their ideas, capabilities, and qualifications. In some instances, akin to coping with the loss of a loved one, they may progress through the stages of death and dying: denial, anger, bargaining, depression, and acceptance (Kubler-Ross, 1969). That is, PIs who start off on the defensive, shielding themselves from the pain of a grant "failing," could switch to the offensive, blaming reviewers for not comprehending the inherent beauty and necessity of the proposed project.

"Research administrators are in a prime position to help faculty unpack reviewer comments, assuming that PIs do indeed wish to understand and learn from the written feedback."

Rather than engaging in a "pity party," as a prelude to providing an impartial analysis of the reviewer comments research administrators can ask PIs one simple question: "Do you think that grant reviews are fair and objective or biased and random?" While there is no right or wrong answer to this attitude-based question, experience suggests that about one-third of PIs believe the reviews are fair and objective, one-third believe they are biased and random, and one-third say, "I don't know." Regardless of their response, an opening now exists to discuss the existence

and nature of the hard-to-please grant reviewer. The conversation begins with an examination of the relationship among grant outcomes, processes, and structures.

The structure-process-outcome analytic model, initially proposed by Avedis Donabedian in the 1960s to assess healthcare quality, is a valuable framework that can be adapted to grantseeking. Modifying Donabedian's definitions (1966) slightly to meet the needs of research administrators, "structure" refers to the environment in which grants are reviewed, "process" refers to the method by which grants are reviewed, and "outcome" refers to the consequence of the grant review. More significantly, these concepts are linked: good structures promote good processes which in turn promote good outcomes (Donabedian, 1988). Thus, when PIs feel a final funding decision—a grant outcome—is unfair, it is valuable to examine the grant structures and processes that led up to the result. The intent is not to find fault or to assign blame—for instance, to the sponsor for having an inadequate grant system, to the reviewer for having insufficient capacity to understand the project, or to the PI for not clearly communicating a project vision. Rather, the aim is to gain insights that will propel the PI forward—to revise and resubmit, to pursue a different sponsor, or to consider a new line of investigation.

". . . the aim is to gain insights that will propel the PI forward—to revise and resubmit, to pursue a different sponsor, or to consider a new line of investigation."

ANALYTIC FRAMEWORK APPLIED: NIH, NSF, ED

To illustrate the structure-process-outcome framework in action, this analytic model is applied to grant programs offered at three federal agencies: the National Institutes of Health (NIH), the National Science Foundation (NSF), and the U.S. Department of Education (ED). Specifically, fourteen dimensions of grant structures and processes, as they existed in FY 2010, are compared and contrasted across five grant programs, including[1]:

- **NIH's Academic Research Enhancement Award (AREA) Parent R15**: supports small research projects in the biomedical and behavioral sciences conducted by students and faculty in health professional schools and other academic components that have not been major recipients of NIH research grant funds.
- **NSF's Transforming Undergraduate Education in Science, Technology, Engineering, and Mathematics (TUES) program**: supports efforts to create, adapt, and disseminate new learning

materials and teaching strategies to reflect advances both in STEM disciplines and in what is known about teaching and learning. The program supports projects representing different stages of development, ranging from small, exploratory investigations to large, comprehensive projects.

- **ED's Fund for the Improvement of Postsecondary Education (FIPSE) Comprehensive Program**: supports and disseminates innovative reform projects that promise to be models for improving the quality of postsecondary education and increasing student access.
- **ED's Hispanic-Serving Institutions (HSI) program**: provides grants to assist HSIs to expand educational opportunities for, and improve the attainment of, Hispanic students. The HSI program grants also enable HSIs to expand and enhance their academic offerings, program quality, and institutional stability.
- **ED's Transition and Postsecondary Programs for Students with Intellectual Disabilities (TPSID) program**: provides grants to institutions of higher education or consortia of institutions of higher education to enable them to create or expand high quality, inclusive model comprehensive transition and postsecondary programs for students with intellectual disabilities.

". . . this article focuses on grant programs that have broad appeal to institutions of higher education, particularly comprehensive universities, liberal arts colleges and community colleges, whose missions center on teaching and learning."

Much has been written already about the ins and outs of securing major research project grants from NIH and NSF; thus, this article focuses on grant programs that have broad appeal to institutions of higher education, particularly comprehensive universities, liberal arts colleges, and community colleges, whose missions center on teaching and learning.

Grant Structures

NIH, NSF, and ED each have grant structures in place that have an effect on grant outcomes. The following seven questions consider dimensions of the environment setting before, during, and after grant applications are reviewed:

- *Before:* How are grant reviewers identified and selected?
- *Before:* How much training do grant reviewers receive?
- *During:* What review criteria are used to evaluate proposals?
- *During:* What scoring rubrics do grant reviewers follow?

- *During:* How do grant reviewers determine their final scores?
- *After:* How are grant reviewers compensated for their efforts?
- *After:* What is the relationship between reviewer scores and funding awards?

Table 1 summarizes responses to these questions and makes clear that similarities and differences exist across programs at these three federal agencies. For instance, on the front side, the sponsors share a common approach to identifying and selecting grant reviewers—namely, a combination of proactive and reactive strategies is used to recruit reviewers with general and specialized experience. Program officers may draw from their knowledge of experts in the field, references included in the proposal, recent professional society programs, literature reviews, volunteers who expressed interest and availability to serve, past grant winners, and, in the case of NSF, recommendations of specific individuals who should and should not review a proposal (Feldman, Meszaros, & Nader, 2007). NSF estimates that over a three-year period, more than half of the PIs who submitted proposals also served as grant reviewers (National Science Foundation, 2008a). What's more, these three sponsors provide some basic training to grant reviewers. Regardless of whether individuals are novice or veteran reviewers, for continuity and consistency purposes all reviewers are usually required to go through the one-hour training session, which includes an overview of the program's purpose and eligible project activities and expenses, an outline of roles and responsibilities and the review schedule, an introduction to the online grants management system, and a general question-and-answer period. The training helps to set context, parameters, and expectations for the review.

On the back side, NIH, NSF, and ED hold a mutual belief that reviewers deserve modest compensation for their volunteer efforts. Whether calculated on a per proposal, per day, or flat rate basis, honoraria for completing a program review are currently in the range of $1,000–$2,000. Given the magnitude and intensity of the task, this stipend often equates to working for nearly three straight weeks at the federal minimum wage rate. No one does the work for the honorarium! They do it for other reasons, such as gaining access to insider information, networking with program officers and professional colleagues, keeping abreast of the latest approaches and techniques, and giving back to the profession (National Institutes of Health, 2011b; National Science Foundation, 2008b; U.S. Department of Education, 2010a). NIH, NSF, and ED also share a common approach to determining funding awards; namely, feedback from reviewers is

advisory only. Program officers analyze the feedback and make recommendations to a higher authority that is ultimately responsible for making the funding award determination. Of the three federal agencies, NSF allows its program officers greater flexibility than NIH and ED in making award recommendations that run counter to feedback from reviewers: an analysis of average reviewer ratings for NSF awards and declines for FY2010 revealed that 1,312 proposals scoring "excellent" were *declined* and 98 proposals scoring "fair to good" were *awarded* (Strausser, 2011). That is to say, higher-scoring proposals can get bumped in favor of lower-scoring proposals that address special considerations, such as geographic distribution of awards, PI status as a novice applicant, organizational status as a primarily undergraduate institution, prior funding history with the sponsor, and extensive collaborative relationships (Miner & Miner, 2008; Miner, Miner, & Griffith, 2011).

". . . higher scoring proposals can get bumped in favor of lower scoring proposals that address special considerations. . . ."

More differences begin to emerge among NIH, NSF, and ED grant structures during the time in which applications are reviewed. While reviewers are all trained to score proposals on their own merit rather than comparing applications, the sponsors take varying approaches to the use of review criteria. For instance, NIH and NSF use standardized review criteria that apply across the vast majority of their programs, including AREA and TUES, respectively, whereas ED review criteria are program-specific. Further, even though the FIPSE, HSI, and TPSID programs are all administratively housed within the ED Office of Postsecondary Education, they utilize assorted review criteria: HSI and TPSID consider "quality of key personnel," but not FIPSE; TPSID and FIPSE consider "significance," but not HSI; they all consider "quality of project evaluation," but only HSI considers "quality of budget."

NIH, NSF, and ED provide reviewers with a scoring rubric, but these tools differ in multiple respects. NIH reviewers must score each review criterion on a 1- to 9-point scale from "exceptional" to "poor"; the final score reflects an assessment of overall impact, not a numerical average of the five review criteria scores. Overall impact expresses a reviewer's estimation of "the likelihood for a project to exert a sustained, powerful influence on the research field(s) involved," but it is not a review criterion in itself (National Institutes of Health, 2011c). Of note, AREA grants (R15), unlike traditional research project grants (R01), do not receive a percentile rank that reflects the approximate percentage of applications

which received better overall impact and priority scores during the past year. At NSF, grant reviewers do not assign numerical ratings to proposals. They consider two standardized review criteria, which need not be weighted equally, in a broad context so that the final rating reflects an overall assessment on a 5-point Likert-type scale from "excellent" to "poor." ED reviewers assign points to each review criterion and then add up all the points to determine a final score; a 100 point scale is used, though TPSID allows bonus points to be earned for meeting competitive priorities—thus, a final score could exceed 100 points. FIPSE weighs review criteria uniformly at 20 points each, whereas HSI and TPSID weigh review criteria independently, with items having values of 7 to 25 points each. What's more, HSI defines point ranges within each criterion. For instance, the "quality of project evaluation" section is worth a total of 15 points and can be assigned points according to the extent to which data elements and data collection procedures are identified: excellent extent = 15–12 points; good extent = 11–9 points; average extent = 8–6 points; minimal extent = 5–1 points; not addressed = 0 points. In other words, HSI minimizes disparities between reviewers who are "tough graders" and "easy graders" because the value of one point is predetermined. FIPSE and TPSID, on the other hand, allow reviewers to settle on their own understanding of the worth of a single point.

Table 1. Grant Structures at NIH, NSF, and ED

	NIH AREA	NSF TUES	ED FIPSE	ED HSI	ED TPSID
Reviewer Selection	Based on general and specialized knowledge, experience, skills	Based on general and specialized knowledge, experience, skills	Based on general and specialized knowledge, experience, skills	Based on general and specialized knowledge, experience, skills	Based on general and specialized knowledge, experience, skills
Reviewer Training	1 hour orientation	1 hour orientation	1 hour orientation	1 hour orientation	1 hour orientation
Review Criteria	• Significance • Investigator • Innovation • Approach • Environment	• Intellectual Merit • Broader Impacts	• Need • Significance • Design • Evaluation • Resources	• Comprehensive Development Plan • Activity Objective • Implementation Strategy • Personnel • Management Plan • Evaluation • Budget	• Need • Significance • Design • Services • Personnel • Resources • Evaluation

Table 1 cont'd.

	NIH AREA	NSF TUES	ED FIPSE	ED HSI	ED TPSID
Scoring Rubric	Each criterion receives a separate score: 1=Exceptional 2=Outstanding 3=Excellent 4=Very Good 5=Good 6=Satisfactory 7=Fair 8=Marginal 9=Poor	Criteria are considered broadly and need not be weighted equally: ● Excellent ● Very Good ● Good ● Fair ● Poor	Criteria are weighted equally; 20 points each	Each criterion is weighted independently, from 7–25 points; within each criterion, point ranges are defined for responses that are excellent, good, average, minimal, and not addressed	Each criterion is weighted independently, from 10–20 points
Final Score Determination	Priority score reflects an assessment of the overall impact, not a numerical average of criteria scores	Rating reflects an overall evaluation	Criteria scores are added together for a maximum score of 100 points	Criteria scores are added together for a maximum score of 100 points	Criteria scores are added together for a maximum score of 100 points; up to 9 additional points may be awarded for meeting competitive priorities
Compensation	$200 per day	$1,200 flat rate	$100 per proposal	$1,000 flat rate	$1,000 flat rate
Award Recommendation	Program officers study established paylines, examine priority scores, analyze feedback from reviewers, and make award recommendations to the Advisory Board/Council; Institute and Center Directors make final funding decisions	Program officers analyze feedback from reviewers and make award recommendations to the Division Director; Division Director makes final funding decision	Program officers analyze feedback from reviewers and make award recommendations to the Secretary; Secretary makes final funding decision	Program officers analyze feedback from reviewers and make award recommendations to the Secretary; Secretary makes final funding decision	Program officers analyze feedback from reviewers and make award recommendations to the Secretary; Secretary makes final funding decision

Grant Processes

NIH, NSF, and ED each have grants processes in place that have an effect on grant outcomes. The following seven questions consider dimensions of the review methodology:

- How many grant reviewers are assigned to each proposal?
- How many proposals are grant reviewers assigned to evaluate?

- What is the maximum length of the proposal's core narrative?
- What type of review is used?
- How much time do grant reviewers have to complete the review?
- What is the average length of grant reviewers' written commentary?
- How are score variances across grant reviewers addressed?

Table 2 summarizes responses to these questions and highlights contrasting approaches used to assess grant proposals submitted to programs at these three federal agencies. These sponsors share commonalities in a couple of aspects of their grants processes. For instance, a minimum of three reviewers evaluate each grant proposal. At ED, only three reviewers score each proposal. At NIH and NSF, review panels may consist of more than three reviewers, perhaps up to twenty reviewers, depending on the degree of specialization needed to assess the proposal. As illustrated in an NIH video (2010) and NSF webinar (Fang & Millard, 2010) of a mock review session, while at least three reviewers will have read each proposal and prepared written comments for examination, other panelists may participate in the general discussion, even if they have not been assigned to read the proposal. NIH, NSF, and ED also permit program officers to take an active role in ensuring that all proposals receive equitable treatment. Program officers may ask questions of reviewers regarding written comments to make certain they are reflective and supportive of the assigned score, as well as make certain they are analytical, in no ways offensive, and grammatically correct.

Differences in grants processes at NIH, NSF, and ED influence the workload of proposal reviewers. At NIH, AREA reviewers had six weeks to review 6–10 proposals that, at their narrative core, were 12 single-spaced pages in length each. At NSF, TUES reviewers had three weeks to review 12 proposals that were 15 single-spaced pages in length each. At ED, FIPSE reviewers had two weeks to review 8–12 proposals that were up to 20 double-spaced pages in length; HSI reviewers had three weeks to review 10 proposals that were 50 double-spaced pages for individual projects or 70 double-spaced pages for cooperative projects; and TPSID reviewers had two weeks to review 28 proposals that were 40 double-spaced pages. In practical terms, TPSID reviewers on average engaged in double the amount of work experienced by HSI, FIPSE, TUES, and AREA reviewers; they needed to complete two reviews on average per day to finish on schedule, whereas the others needed to complete one review or less on average per day.

NIH, NSF, and ED each employed, at a minimum, a mail review. Proposals were made available electronically to reviewers (i.e., the eco-friendly version of proposals

being copied and mailed) and were reviewed independently from the comforts of their homes and offices. A reviewer's evaluation form was completed online for each proposal, which included an assessment score and approximately 1–4 single-spaced pages of analysis commentary. Of the five programs, FIPSE was the only one to use a mail review exclusively. Said differently, FIPSE reviewers did not have opportunity to meet each other or to share their perceptions of proposal merits. On the other hand, AREA and TUES reviewers engaged in mail and in-person panel reviews. Reviewers critiqued proposals individually and uploaded their analysis into an online grants management system. Subsequently, they convened in Washington, DC for 2–3 days to conduct panel discussions in person, modify comments, and make final submissions. AREA reviewers knew that proposals ranking in the bottom half of all applications would not be discussed or receive priority scores and TUES reviewers knew that every application would be discussed and scored. Nevertheless, completing the reviewer evaluation forms online in advance of the panel review encouraged reviewers to prepare for deliberations. HSI and TPSID reviewers, in addition to mail and in-person panel reviews, participated in teleconference panel reviews. That is, after reviewers completed their independent assessments of proposals and uploaded their responses into an online grants management system, program officers coordinated conference calls, typically one hour in length, where proposal strengths, weaknesses, and scores were discussed. During the final week of the HSI and TPSID review process, reviewers came together in Washington, DC to panel the remaining proposals in person. The advantage of the teleconference panel reviews is that they allowed chemistry to develop among reviewers, which fostered collegial dialogue during the in-person panel reviews.

"In the words of the program officer, 'Institutions are notorious about challenging reviewers' comments when there is a large discrepancy—and, of course they give more credibility to the higher score than the lower score as being justified!'"

The NIH, NSF, and ED programs have their own procedures for managing instances of significant variance among reviewers' scores. AREA and TUES have reviewers upload their individual critiques in advance of the in-person panel review. As a result, program officers have an early alert to which proposals may require substantive discussion. Program officers are allowed to ask reviewers to further justify their comments if an inconsistency seems to

exist with the rating; reviewers are not obligated to resolve differences in scoring. For AREA proposals, after paneling, individual reviewer scores are averaged and the result is multiplied by 10 to determine the final priority score. For TUES proposals, based on the panel discussion, a collective summary review is written; however, rather than averaging scores, individual ratings are retained to provide an indication of the relative merits of different perspectives. FIPSE is a unique case precisely because reviewers are only allowed to have discussions with program officers, not other reviewers. During the FIPSE training program, reviewers are informed that when a variance of 15 points or more exists, they will be contacted and given the opportunity to reconsider their scores and comments; reviewers are not obligated to change their scores, but further justification may be necessary. In the words of the program officer, "Institutions are notorious about challenging reviewers' comments when there is a large discrepancy—and, of course they give more credibility to the higher score than the lower score as being justified!" HSI and TPSID tolerate a more narrow range of variance: applications with a 10-point or more divergence in scores must be discussed during the panel review (U.S. Department of Education, 2010b, 2010c). If reviewers are not able to reach closer agreement after substantial dialogue, they must complete and sign a "Record of Discussion" form that details the basis of the impasse. To avoid making additional work for themselves, reviewers are frequently willing to compromise to keep scores within the acceptable variance range. Put another way, the hard-to-please reviewer is more difficult to spot at ED than at NIH and NSF.

IMPLICATIONS

Undeniably, receiving a sponsor's "I regret to inform you that…" negative funding decision letter is disappointing. More aggravating, though, are the reviews that contain mixed, or even contradictory, feedback. In a survey of unsuccessful grant applicants, Wood, Meek and Harman (1992) identified five specific aspects of inconsistencies in ratings between grant reviewers that were of greatest concern to PIs: (a) a mismatch of topic and reviewers; (b) a lack of understanding of the topic; (c) a perceived bias against high-risk research; (d) the brevity of comments provided; and (e) a failure to substantiate criticisms. In the following discussion, each of these concerns is examined through the lens of the structure-process-outcome analytic model. Also offered are strategies research administrators can use to help PIs understand the existence and nature of the hard-to-please grant reviewer. As a result, PIs will be better positioned to determine appropriate next steps.

Table 2. Grant Processes at NIH, NSF, and ED

	NIH AREA	NSF TUES	ED FIPSE	ED HSI	ED TPSID
Number of Reviewers	3+	3+	3	3	3
Number of Proposals	6–10	12	8µ12	10	28
Proposal Length	12 SS page	15 SS page	20 DS page	50 DS page, 70 DS page	40 DS page
Review Type	• Mail • In-person panel	• Mail • In-person panel	• Mail	• Mail • Teleconference • In-person panel	• Mail • Teleconference • In-person panel
Time to Review	6 weeks	3 weeks	2 weeks	3 weeks	2 weeks
Response Length	1–3 pages	1–2 pages	1–2 pages	1–4 pages	1–3 pages
Score Variance	Program officers may ask questions during panel; individual reviewer scores are averaged and the result multiplied by 10 to determine the priority score	Program officers may ask questions during panel; they may ask for comments to be further justified if an inconsistency seems to exist with the rating	Proposals with a 15 point or more divergence in scores will serve as a prompt for program officers to ask for comments to be further justified or for scores to be reconsidered	Proposals with a 10 point or more divergence in scores must be discussed during panel; a record of discussion must be completed when no closer agreement is reached	Proposals with a 10 point or more divergence in scores must be discussed during panel; a record of discussion must be completed when no closer agreement is reached

A Mismatch of Topic and Reviewer

NIH, NSF, and ED have established grants structures for identifying and selecting reviewers to evaluate the applications they receive. These sponsors typically assemble a group of individuals who possess specialized knowledge of the topic and general knowledge of the field. Whether participating in one type or a combination of mail, in-person panel, and teleconference panel reviews, reviewers provide feedback that, collectively, assesses the specifics of the project and the larger context in which it is situated (Jayasinghe, March, & Bond, 2001; Trumbo, 1989). Sponsor efforts to balance the representation of a depth and breadth of experience could contribute to PIs' perceptions that a mismatch exists between proposal topics and reviewers. In these instances, PIs may be assuming that hard-to-please reviewers are (poorly) filling the role of subject matter specialists when in reality they are filling the role of branch generalists.

". . . because there is considerable turnover annually of individual reviewers and of review panels, it can pay to be persistent; funding chances usually improve with resubmissions."

In other words, elements of chance exist within grants structures that influence grant outcomes; namely, the assignment of specific proposals to specific reviewers and the assignment of individual reviewers to in-person and/or teleconference panels. Cole, Cole, and Simon (1981) offered a quantifiable estimate of the impact of this phenomena: "The fate of a particular grant application is roughly half determined by the characteristics of the proposal and the principal investigator, and about half by apparently random elements which might be characterized as the 'luck of the reviewer draw'" (p. 885). Lead reviewers play a critical role in setting the tone for the rest of the panel. While differences in professional opinion are inevitable among reviewers, the manner in which they are handled can affect the final result. One veteran NIH grant reviewer reported, "I have been in a couple of meetings where people have felt their reputation was at stake, and it was a personal insult if the committee didn't agree with them. And I've been on others in disagreement where the tone was, instead, 'Oh I'm so sorry. I must have missed something,' as opposed to 'You're wrong and I'll prove it to you.' You can just feel the difference" (Hebert, 2002). More significantly for PIs, because there is considerable turnover annually in individual reviewers and in review panels, it can pay to be persistent; funding chances usually improve with resubmissions. Consider: in FY2010, NIH research project grants that were new submissions had an 11.5% success rate whereas resubmissions had a 34.9% success rate (National Institutes of Health, 2011d). Research administrators should encourage PIs who are turned down the first time to consider revising and resubmitting the proposal based on feedback from reviewers (Miner & Miner, 2008).

A Lack of Understanding of the Topic

Even when NIH, NSF, and ED grants structures suitably match proposal topics and reviewers, PIs may hold tight to the belief that the hard-to-please reviewer is not expert *enough* to fully grasp the subject matter. In practicality, in any collection of three or more reviewers, including ones with similar academic credentials, differences will still exist in their experiences, ideas, and ways of understanding the world. Consequently, there is intersubjectivity among reviewers. Three reviewers may share the same point-of-view on a proposal and yet score it in a different way (Cole, Cole, & Simon, 1981). At NIH, this manifests itself when reviewer #1's rating of "outstanding" equates to

reviewer #2's rating of "excellent," which equates to reviewer #3's rating of "very good." All three scores are respectable, yet PIs perceive reviewer #3 as being hard-to-please because the score reflects the proposal as having a medium impact while the other two scores reflect it as having a high impact.

In an interview with faculty who were veteran NIH, NSF, and U.S. Department of Agriculture grant reviewers as well as grant writers, one offered the following sage advice: "I used to write to a peer; now I write to a committee. I write to teach both the specialist scholar in my particular field and the generalists, who make up the majority of the panel" (Porter, 2005, p. 9). This notion of grant writer-as-teacher is particularly important not only when the specialist-generalist gap needs bridging but also when reviewers share the same field as the PI but not the same sub-field. Persuading these hard-to-please reviewers means working from the known to the unknown. Research administrators can probe with PIs to target an appropriate level of shared understanding from which a case can be built; PIs can subsequently adapt their narrative to meet reviewers' needs and expectations (Ede & Lunsford, 1984). When done effectively, the proposal will motivate one or more reviewers to act as a "champion" for the project (Altman, 2009; Obrecht, Tibelius, & D'Aloisio, 2007). Inspired reviewers will, for instance, read articles cited in the bibliography to gain a greater familiarity with the topic (Member, 2003; Molfese, Cervelin, & Miller, 2007; Trumbo, 1989). Thus, reviewers can advocate on the PI's behalf for a favorable enough final score that the project will garner grant funding.

A Perceived Bias against High-Risk Research

Beyond the identification and selection of reviewers, NIH, NSF, and ED have grant structures in place to train grant reviewers. Training sessions typically last about one hour and address contextual, logistical, and technological considerations. For example, reviewers are briefed on the sponsor's mission and the priorities of the grant program, their performance expectations and the timeline for the review, and nuances of the online grants management system. In their roles, reviewers act as the conscience of the community, ensuring that grant funds are spent wisely. To catch the attention of reviewers, PIs recognize that, among a stack of proposals, projects must have some intellectual sex appeal. Yet some PIs also hold the perception that a bias exists against high-risk research (Berezin, 1998; Obrecht, Tibelius, & D'Aloisio, 2007; Wood, Meek, & Harman, 1992). Unless they have served as grant reviewers themselves, these PIs may not fully appreciate that being a good steward of sponsor funds often means distinguishing scientific fact from professional judgment and may mean

supporting research that takes the next methodical step along established lines (Office of Management and Budget, 2004).

This perceived bias against high-risk research is particularly sensitive at NIH. "Innovation" is one of five review criteria NIH uses to assess grant applications, and a series of questions guide reviewers in their consideration of the extent to which a project is path-breaking:

> Does the application challenge and seek to shift current research or clinical practice paradigms by utilizing novel theoretical concepts, approaches or methodologies, instrumentation, or interventions? Are the concepts, approaches or methodologies, instrumentation, or interventions novel to one field of research or novel in a broad sense? Is a refinement, improvement, or new application of theoretical concepts, approaches or methodologies, instrumentation, or interventions proposed? (National Institutes of Health, 2011c)

In her "Rock Talk" blog, Sally Rockey, NIH Deputy Director for Extramural Research, presented the results of a study of the correlation between overall impact scores and criterion scores on nearly 55,000 research grant applications submitted in FY2010. Multiple regression analysis ranked the order in which criteria carried the greatest to least weight: approach, significance, innovation, investigator, and environment. That is to say, reviewers consider innovation as a central but not the foremost among NIH review criteria. With this in mind, research administrators can help PIs to reframe their project so that it is "innovative" but not so "highly innovative" that it will be deemed "risky." Research administrators can also remind PIs that reviewer comments should not be interpreted as personal attacks. According to Buller (2002), drawing from his experiences as an NIH reviewer, "when a reviewer is interested in an idea, they often provide tough, detailed critiques of an application in order to help the researchers ultimately produce an application that can be evaluated highly by the entire review team" (p. 414). This sentiment is also shared by Wiley in his aptly titled article, "Peer review isn't perfect…But it's not a conspiracy designed to maintain the status quo" (2008).

The Brevity of Comments Provided

One aspect of the reviewer training provided by NIH, NSF, and ED, respectively, deals with the preparation of written comments. They emphasize that grant reviews are written for two audiences—program staff and applicants. Program staff use the feedback to justify award recommendations and applicants rely on feedback to determine whether a declined proposal should be revised and

resubmitted. The sponsors offer a range of supports to reviewers: questions to consider associated with each review criterion, checklists of review do's and don'ts, samples of completed reviewer evaluation forms, and names of program staff who can be contacted to clarify technical and financial aspects of the review. ED even specifies minimum length thresholds for written comments:

> Both successful and unsuccessful applicants will appreciate receiving Technical Review Forms that correspond with their applications with three or more lines of comprehensive, written comments. **ONE WORD/ONE LINE WRITTEN COMMENTS FOR STRENGTHS OR WEAKNESSES WILL NOT BE ACCEPTED** (U.S. Department of Education, 2010c, p. 10, *emphasis original*).

While these grant structures set performance expectations for reviewers, sponsors' grant processes influence the fidelity to which they are followed. It often boils down to a workload issue (Abrams, 1991; Cook et al., 2005). Reviewers attempt to balance the quality of feedback with the number and length of proposals and the time available to complete the review. Likewise during the review period, program officers attempt to balance the number of requests they make of reviewers to edit the previous reviewer evaluation form with the need to press forward and complete the next proposal review.

"Program staff use the feedback to justify award recommendations and applicants rely on feedback to determine whether a declined proposal should be revised and resubmitted."

Repeated calls for greater substance and clarity in commentary can come at the expense of maintaining an overall time-to-review equity and of adhering to the established review timeline. ED offers a conservative estimate that reviewers will devote two to three hours per application, including time to read the proposal, write an evaluation, and submit the review online; however, ED's online grants management system, G5, is set to time out after 20 minutes; thus, reviews are often composed in a separate word-processing program and then responses for each criterion are cut-and-pasted into G5. This extra step increases the time spent on each application and the overall workload, but does not enhance the volume or quality of feedback. Regardless of the length of comments, research administrators should encourage PIs to contact program officers to discuss reviews. In response to an open-ended request such as, "Help me to

understand the apparently divergent perspectives reviewers had on my proposal," program officers often provide reflections and advice over the phone that is much more candid than they would ever put in an email message.

A Failure to Substantiate Criticisms

Though NIH, NSF, and ED each provide relatively standardized orientation trainings, reviewers internalize and apply this information in different ways. In-person and teleconference panels, especially, tend to develop their own chemistry (Obrecht, Tibelius, & D'Aloisio, 2007; Porter, 2009; Tufts University, 2010). Reviewers find their expertise niche—specialized content knowledge, familiarity with instrumentation, general experience with target populations, knowledge of pedagogy, skill in assessment, proficiency with budgets—as they dialogue with one another. According to research by Klahr (1985), who served as an NSF panel review member for several years, "the positive aspects of a proposed study are usually acknowledged by all of its reviewers, but different weaknesses are discovered by different panelists" (p. 150). Individual written critiques, then, may demonstrate varying depths of analysis: a reviewer drawing from personal grounding can spell out criticisms in detail whereas other panelists, who agree at least to some extent with the reviewer but are outside of their expertise niche, are left to capture the spirit of the concern and summarize it in their own words. Further, the grant processes at ED, more so than at NIH and NSF, strongly encourage reviewers to reconcile variances in proposal scores. ED reviewers discuss, negotiate, and horse-trade to get scores within the acceptable range so that they do not have to write a dissenting opinion. Because the final score is an average of the sum of panelist's criteria scores, savvy reviewers know where they can compromise on points to get within a range at the criteria level and yet not change the overall funding recommendation. Consider: prior to paneling, reviewers might have initial scores of 91, 86, and 75, which produce an average score of 84 and have a variance of 16 points. After deliberation, reviewers might adjust their scores to 88, 84, and 80, which also produce an average score of 84 but have a variance of only 8 points. As a consequence of this attention to scoring, individual critiques may be thinly written. When viewed together, however, reviewers' written commentary presents an air of authority with the collective wisdom and judgment of the group.

Beyond examining intra-panel agreement, researchers have also investigated the extent to which inter-panel agreement exists. To be precise, two sets of panels were given the same proposals to review independently and their funding recommendations were compared. Johnson (2008) conducted a statistical analysis of R01

grant applications at NIH and concluded that 25% of funding decisions would be reversed had the proposals been reviewed by a different panel. Cole, Cole, and Simon (1981) examined grant applications submitted to three programs at NSF and estimated a reversal rate of 24–30%. Obrecht, Tibelius, and D'Aloisio (2007) studied research training fellowship proposals submitted to the Canadian Institute of Health Research and determined that 27% of applications would be overturned by a new panel of reviewers. It is logical that some reversals would occur near the payline. That is, Panel A might rank a proposal 9th while Panel B might rank a proposal 11th and the cutoff for funding was the 10th proposal; the funding decision would be reversed by the second panel. Research administrators can urge PIs to solicit the program officer's reactions to the written comments, namely, to find out which of the weaknesses cited were the most significant and how far below the funding line the proposal landed. A PI might find the proposal was a near-miss and that funding success is really closer than originally imagined.

CONCLUSION

It may be natural for a PI, frustrated in general by a negative funding decision and specifically by diametrically opposed reviewer comments, to question whether grant reviews are more biased and random than fair and objective. However, among faculty who have served as grant reviewers, there is a firm belief in the objectivity of review panels. While acknowledging that peer review is not perfect, they cite as strengths the transparency of the grant system (Kessel, 2006), explicit and uniform processes used (Wiley, 2008), and democratic and self-correcting quality of panels (Porter, 2005), and they underscore that instances of bias and cronyism are infrequent (Molfese, Cervelin, & Miller, 2007). These faculty also point out that volunteering to serve as a reviewer is a great way for PIs to learn the inner workings of grant programs and panels. To a similar end, when the structure-process-outcome analytic model is applied to NIH, NSF, and ED, it can be seen that hard-to-please reviewers exist in programs at all three federal agencies. Given their respective grant structures and processes, though, hard-to-please reviewers are more difficult to pinpoint at ED because they are masked through the formula for reconciling score variance. An element of random chance is also present in the grant system of each of these sponsors—chiefly, the assignment of specific proposals to specific reviewers and the assignment of individual reviewers to review panels.

"With the benefit of time and some distance, PIs may recognize that reviewer commentary contains kernels of truth—sections weren't as clear as they could have been, details were omitted, alternative options were not considered—that can actually lead to a stronger application."

More broadly, research administrators can use the structure-process-outcome framework with PIs to contextualize reviewer feedback and to position them for future grant success. Research administrators should encourage PIs to always contact program officers for a debriefing after learning a grant outcome. When a proposal is funded, PIs can talk with program officers to determine exactly which dimensions of the project caught reviewers' attention and whether there are minor issues that need to be addressed; when a proposal is not funded, PIs can find out which aspects were of the greatest concern to reviewers and what it might take to convince them that the project does indeed merit funding. Of note, debriefing is a time to listen, not to argue. Aggressively challenging the program officer and formally appealing the grant outcome are unlikely to be successful. Differences in judgment will not change the final result; only when it can be proven that a procedural mistake occurred in the grant review process do PIs have even the slightest chance of overturning a funding decision (Trumbo, 1989). PIs are better off spending that emotional and intellectual energy systematically analyzing reviewer feedback and then, as appropriate, revising and resubmitting their proposals. Research administrators can counsel PIs to let the reviews sit for a short period and come back to them with fresh eyes, as if written by a friend (Wiley, 2008). With the benefit of time and some distance, PIs may recognize that reviewer commentary contains kernels of truth—sections weren't as clear as they could have been, details were omitted, alternative options were not considered—that can actually lead to a stronger application. Tinkering changes to the narrative, most likely, will not be enough. At the same time, proposals do not need to be perfect to attract funding; rather, they need to be persuasive (Miner & Miner, 2005). PIs who anticipate the needs of reviewers and satisfy those expectations in the proposal increase their odds for a positive grant outcome.

ENDNOTE

1. This bulleted list represents five of the items in the literature cited: National Institutes of Health, 2011a; National Science Foundation, 2011; U.S. Department of Education, 2011a,b,c. The grant programs are described as they existed in FY 2010; specific websites were accessed in 2011.

LITERATURE CITED

Abrams, P. A. (1991). The predictive ability of peer review of grant proposals: The case of ecology and the U.S. National Science Foundation. *Social Studies of Science, 21*(1), 111–132.

Altman, M. (2009). Funding, funding. *PS: Political Science & Politics, 42*(1), 521–526.

Berezin, A. (1998). The perils of centralized research funding systems. *Knowledge, Technology & Policy, 11*(3), 5–26.

Buller, D. B. (2002). Final thoughts on funded communication research. *Journal of Applied Communication Research, 30*(4), 411–417.

Cole, S., Cole, J. R., & Simon, G. A. (1981). Chance and consensus in peer review. *Science, 214*(4523), 881–886.

Cook, W. D., Golany, B., Kress, M, Penn, M., & Raviv, T. (2005). Optimal allocation of proposals to reviewers to facilitate effective ranking. *Management Science, 51*(4), 655–661.

Donabedian, A. M. (1966). Evaluating the quality of medical care. *Milbank Memorial Fund Quarterly, 44*(1), 166–203.

Donabedian, A. M. (1988). The quality of care: How can it be assessed? *Journal of the American Medical Association, 260*(12), 1743–1748

Ede, L., & Lunsford, A. (1984). Audience addressed/audience invoked: The role of audience in composition theory and pedagogy. *College Composition and Communication, 35*(2), 155–171.

Fang, N., & Millard, D. (2010). NSF proposal review process: Mock review webinar. Retrieved October 6, 2011 from http://www.odu.edu/ao/research/proposaldev/workshops.shtml

Feldman, J., Meszaros, J., & Nader, R. (2007). *NSF proposal preparation highlights*. Retrieved June 12, 2011 from http://nuweb.neu.edu/nhe/Proposalprepst.louis07.ppt

Hebert, R. (2002). *Sitting in judgment: Myths and realities of peer review*. Association for Psychological Science. Retrieved October 6, 2011 from http://www.psychologicalscience.org/observer/getArticle.cfm?id=906

Jayasinghe, U. W., Marsh, H. W., & Bond, N. (2001). Peer review in the funding of research in higher education: The Australian experience. *Educational Evaluation and Policy Analysis, 23*(4), 343–364.

Johnson, V. E. (2008). Statistical analysis of the National Institutes of Health peer review system. *PNAS, 105*(32), 11076–11080.

Kessel, D. (2006). Writing successful grant applications for preclinical studies. *CHEST, 130*(1), 296–298.

Klahr, D. (1985). Insiders, outsiders, and efficiency in a National Science Foundation panel. *American Psychologist, 40*(2), 148–154.

Kubler-Ross, E. (1969). *On death and dying*. New York: Macmillan.

Member, P. (2003). NSF grant reviewer tells all. *Science Careers* (online publication of *Science* magazine). Retrieved June 12, 2011 from http://sciencecareers.sciencemag.org/career_development/previous_issues/articles/2310/nsf_grant_reviewer_tells_all/

Miner, J. T., & Miner, L. E. (2005). *Models of proposal planning & writing*. Westport, CT: Praeger.

Miner, J. T., & Miner, L. E. (2008). *Proposal planning & writing, 4th ed.* Westport, CT: Greenwood.

Miner, J. T., Miner, L. E., & Griffith, J. (2011). *Collaborative grantseeking: A guide to designing projects, leading partners, and persuading sponsors*. Santa Barbara, CA: Greenwood.

Molfese, V. J., Cervelin, J., & Miller, P. F. (2007). Demystifying the NIH proposal review process. *Journal of Research Administration, 38*(1), 96–106.

National Institutes of Health. (2010). NIH grant review process YouTube videos. Retrieved October 6, 2011 from http://cms.csr.nih.gov/ResourcesforApplicants/InsidetheNIHGrantReviewProcessVideo.htm

National Institutes of Health. (2011a). *Types of grant programs*. Retrieved June 12, 2011 from http://grants.nih.gov/grants/funding/funding_program.htm

National Institutes of Health. (2011b). *Do you want a rewarding challenge? Consider becoming an NIH reviewer*. Retrieved October 6, 2011 from http://cms.csr.nih.gov/nr/rdonlyres/3ac839de-4845-4269-b4e0-a9900f7d4cbb/26055/reviewfornih.pdf

National Institutes of Health. (2011c). *Peer review process*. Retrieved June 12, 2011 from http://grants.nih.gov/grants/peer_review_process.htm

National Institutes of Health. (2011d). *Success rates*. Retrieved June 12, 2011 from http://report.nih.gov/success_rates/index.aspx

National Science Foundation. (2008a). *Broadening participation at the National Science Foundation: A framework for action*. Retrieved June 12, 2011 from http://www.nsf.gov/od/broadeningparticipation/nsf_frameworkforaction_0808.pdf

National Science Foundation. (2008b). *Why you should volunteer to serve as an NSF reviewer*. Retrieved October 6, 2011 from http://www.nsf.gov/bfa/dias/policy/meritreview/reviewer.jsp

National Science Foundation. (2011). *Transforming undergraduate education in science, technology, engineering, and mathematics*. Retrieved June 12, 2011 from http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5741

Obrecht, M., Tibelius, K., & D'Aloisio, G. (2007). Examining the value added by committee discussion in the review of applications for research awards. *Research Evaluation, 16*(2), 79–91.

Office of Management and Budget. (2004). *Final information quality bulletin for peer review*. Retrieved June 12, 2011 from http://www.whitehouse.gov/sites/default/files/omb/assets/omb/memoranda/fy2005/m05-03.pdf.

Porter, R. (2005). What do grant reviewers really want, anyway? *Journal of Research Administration, 36*(2), 5–13.

Porter, R. (2009). Can We Talk? Contacting grant program officers. *Research Management Review, 17*(1), 10–16.

Rockey, S. (2011). *Correlation between overall impact scores and criterion scores*. Retrieved June 12, 2011 from http://nexus.od.nih.gov/all/2011/03/08/overall-impact-and-criterion-scores/

Strausser, B. (2011). NSF update: NCURA Region IV Spring Meeting, April 17–20, 2011. Retrieved October 6, 2011 from http://www.ncuraregioniv.com/meetings/mtg_info/2011%20OH%20Meeting/NSF%20Update%20Spring%202011_bas.pdf

Trumbo, B. E. (1989). How to get your first research grant. *Statistical Science, 4*(2), 121–131.

Tufts University, Office of Proposal Development. (2010). *What we learned on our summer vacation: NIH Reviewers tell us how the 12-page R01 does (and doesn't) change things*. Retrieved June 12, 2011 from http://grantwriting.tufts.edu/downloads/NIH_Reviewer_Focus_Groups.pdf

U.S. Department of Education. (2010a). *Fund for the Improvement of Postsecondary Education, Comprehensive Program (CFDA 84.116B) Reviewer's guide*. Retrieved June 12, 2011 from http://reviewgrants.com/fipse/documents/Reviewers_Guide_2010_Comprehensive.pdf

U.S. Department of Education. (2010b). Hispanic-Serving Institutions Program (CFDA 84.031S) Field Review Orientation. Retrieved June 12, 2011 from http://reviewgrants.com/hsi/documents/2010_HSI_Orientation.pdf

U.S. Department of Education. (2010c). *Model comprehensive transition and postsecondary programs for students with intellectual disabilities* (CFDA 84.407A) Instructional Handbook for Reviewers. Retrieved June 12, 2011 from http://reviewgrants.com/tpsid/documents/2010%20TPSID%20Reviewer's%20Handbook3.pdf

U.S. Department of Education. (2011a). *Fund for the Improvement of Postsecondary Education-comprehensive program*. Retrieved June 12, 2011 from http://www2.ed.gov/programs/fipsecomp/index.html

U.S. Department of Education. (2011b). *Developing Hispanic-Serving Institutions Programs—Title V*. Retrieved June 12, 2011 from http://www2.ed.gov/programs/idueshsi/index.html

U.S. Department of Education. (2011c). *Transition and postsecondary programs for students with intellectual disabilities*. Retrieved June 12, 2011 from http://www2.ed.gov/programs/tpsid/index.html

Wiley, S. (2008). Peer review isn't perfect…But it's not a conspiracy designed to maintain the status quo. *Scientist, 22*(11), 31.

Wood, F. Q., Meek, V. L., & Harman, G. (1992). The research grant application process: Learning from failure? *Higher Education, 24*(1), 1–23.